GenCore version 4.5

OM protein - protein search, using sw model

Run on: March 1, 2001, 16:09:23 ; Search time 299.73 Seconds
(without alignments)
26.982 Million cell updates/sec

Title: US-09-331-631A-5_COPY_76_144
Perfect score: 381
Sequence: 1 NRQRDPQQQYEQCQKRCQRR.....EEQQREDEEKYEERMKEGDN 69

Scoring table: BLOSUM62
Gapop 10.0 , Gapext 0.5

Searched: 374700 seqs, 117207915 residues

Total number of hits satisfying chosen parameters: 374700

Minimum DB seq length: 0
Maximum DB seq length: 2000000000

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : SPTREMBL_15:*
1: sp_archea:*
2: sp_bacteria:*
3: sp_fungi:*
4: sp_human:*
5: sp_invertebrate:*
6: sp_mammal:*
7: sp_mhc:*
8: sp_organelle:*
9: sp_phage:*
10: sp_plant:*
11: sp_rodent:*
12: sp_virus:*
13: sp_vertebrate:*
14: sp_unclassified:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	381	100.0	625	10	Q9SPL3	Q9spl3 macadamia i
2	363	95.3	666	10	Q9SPL5	Q9spl5 macadamia i
3	359	94.2	666	10	Q9SPL4	Q9spl4 macadamia i
4	173	45.4	525	10	Q43358	Q43358 theobroma c
5	166	43.6	593	10	Q9SEW4	Q9sew4 juglans reg
6	115	30.2	810	10	Q9ZWI3	Q9zwi3 cucurbita m
7	114.5	30.1	1038	5	O60983	O60983 dictyosteli
8	109	28.6	554	5	Q9VPS3	Q9vps3 drosophila
9	109	28.6	1089	12	O40947	O40947 kaposi's sa
10	105.5	27.7	385	5	Q17909	Q17909 caenorhabdi
11	102	26.8	1162	12	Q98148	Q98148 kaposi's sa
12	101	26.5	1737	5	Q9TW28	Q9tw28 dictyosteli
13	100	26.2	838	5	Q9Y0C9	Q9y0c9 dictyosteli
14	100	26.2	953	4	Q14151	Q14151 homo sapien
15	100	26.2	1129	12	Q9QR71	Q9qr71 kaposi's sa
16	100	26.2	1590	5	P91152	P91152 caenorhabdi
17	99	26.0	1390	5	O77033	O77033 dictyosteli
18	98.5	25.9	2701	4	Q9Y520	Q9y520 homo sapien
19	97.5	25.6	1027	4	Q9NST7	Q9nst7 homo sapien
20	97.5	25.6	1165	4	O95819	O95819 homo sapien
21	97.5	25.6	1175	4	O75172	O75172 homo sapien
22	97.5	25.6	1233	11	P97820	P97820 mus musculu
23	97.5	25.6	1257	4	O95033	O95033 homo sapien
24	97.5	25.6	1700	5	Q9VVA9	Q9vva9 drosophila
25	97	25.5	378	2	O31328	O31328 borrelia bu
26	97	25.5	378	2	O08264	O08264 borrelia bu
27	97	25.5	1268	4	Q9UKD8	Q9ukd8 homo sapien
28	97	25.5	1276	4	Q9UKE1	Q9uke1 homo sapien
29	97	25.5	1323	4	Q9UKE0	Q9uke0 homo sapien
30	97	25.5	1331	4	Q9UKE4	Q9uke4 homo sapien
31	96	25.2	555	5	Q9U0V4	Q9u0v4 leishmania
32	96	25.2	1297	4	Q9UKD9	Q9ukd9 homo sapien
33	96	25.2	1305	4	Q9UKE3	Q9uke3 homo sapien
34	96	25.2	1333	4	O60298	O60298 homo sapien
35	96	25.2	1352	4	Q9UKE2	Q9uke2 homo sapien
36	96	25.2	1360	4	Q9UKE5	Q9uke5 homo sapien
37	95.5	25.1	375	2	O87298	O87298 borrelia bu
38	95.5	25.1	483	10	O49927	O49927 pisum sativ
39	95.5	25.1	1026	5	Q9VLJ3	Q9vlj3 drosophila
40	95.5	25.1	1295	4	Q9P2R8	Q9p2r8 homo sapien
41	95.5	25.1	1300	11	Q9JM92	Q9jm92 mus musculu
42	95.5	25.1	1303	4	Q9P1X1	Q9p1x1 homo sapien
43	95.5	25.1	1308	11	Q9JM52	Q9jm52 mus musculu
44	94.5	24.8	876	11	Q9WU62	Q9wu62 mus musculu
45	94	24.7	233	4	Q9UFP2	Q9ufp2 homo sapien

ALIGNMENTS

RESULT 1
Q9SPL3
ID Q9SPL3 PRELIMINARY; PRT; 625 AA.
AC Q9SPL3;
DT 01-MAY-2000 (TrEMBLrel. 13, Created)
DT 01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT 01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE VICILIN PRECURSOR (FRAGMENT).
GN AMP2.
OS Macadamia integrifolia (Macadamia nut).
OC Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX NCBI_TaxID=60698;
RN [1]
RP SEQUENCE FROM N.A.
RC TISSUE=NUT KERNEL;
RA Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT "A family of antimicrobial peptides is produced by processing of a 7S
RT globulin protein in Macadamia integrifolia.";
RL Plant J. 0:0-0(1999).
DR EMBL; AF161885; AAD54246.1; -.
DR HSSP; P02853; 2PHL.
DR INTERPRO; IPR001113; -.
DR PFAM; PF00546; Seedstore_7s; 1.
FT NON_TER 1 1
SQ SEQUENCE 625 AA; 73586 MW; 415808A89D370296 CRC64;

Query Match 100.0%; Score 381; DB 10; Length 625;
Best Local Similarity 100.0%; Pred. No. 3.1e-27;
Matches 69; Conservative 0; Mismatches 0; Indels 0; Gaps 0;

```
Qy       1 NRQRDPQQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQKRYEEQQREDEEKY 60
           ||||||||||||||||||||||||||||||||||||||||||||||||||||||||||||
Db      76 NRQRDPQQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQKRYEEQQREDEEKY 135

Qy      61 EERMKEGDN 69
           |||||||||
Db     136 EERMKEGDN 144
```

RESULT 2

```
Q9SPL5
ID   Q9SPL5      PRELIMINARY;      PRT;   666 AA.
AC   Q9SPL5;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   VICILIN PRECURSOR.
GN   AMP2.
OS   Macadamia integrifolia (Macadamia nut).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX   NCBI_TaxID=60698;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=NUT KERNEL;
RA   Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT   "A family of antimicrobial peptides is produced by processing of a 7S
RT   globulin protein in Macadamia integrifolia kernels.";
RL   Plant J. 0:0-0(1999).
DR   EMBL; AF161883; AAD54244.1; -.
DR   HSSP; P02853; 2PHL.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
SQ   SEQUENCE   666 AA;  78217 MW;  C752B884B2DF0224 CRC64;


  Query Match             95.3%;  Score 363;  DB 10;  Length 666;
  Best Local Similarity   95.7%;  Pred. No. 1.4e-25;
  Matches   66;  Conservative    0;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 NRQRDPQQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQKRYEEQQREDEEKY 60
           |||||||||||||||| |||||||||||| ||||||||||||||||||||||||||||||
Db     117 NRQRDPQQQYEQCQKHCQRRETEPRHMQTCQQRCERRYEKEKRKQQKRYEEQQREDEEKY 176

Qy      61 EERMKEGDN 69
           ||||||  ||
Db     177 EERMKEEDN 185


RESULT   3
Q9SPL4
ID   Q9SPL4      PRELIMINARY;      PRT;   666 AA.
AC   Q9SPL4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)\
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   VICILIN PRECURSOR.
GN   AMP2.
OS   Macadamia integrifolia (Macadamia nut).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; Proteaceae; Macadamia.
OX   NCBI_TaxID=60698;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=NUT KERNEL;
RA   Marcus J.P., Goulter K.C., Green J.L., Manners J.M.;
RT   "A family of antimicrobial peptides is produced by processing of a 7S
RT   globulin protein in Macadamia integrifolia.";
RL   Plant J. 0:0-0(1999).
DR   EMBL; AF161884; AAD54245.1; -.
DR   HSSP; P02853; 2PHL.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
SQ   SEQUENCE   666 AA;  78243 MW;  0ECA22F8710F8A7B CRC64;


  Query Match             94.2%;  Score 359;  DB 10;  Length 666;
  Best Local Similarity   94.2%;  Pred. No. 3.2e-25;
  Matches   65;  Conservative    1;  Mismatches    3;  Indels    0;  Gaps    0;

Qy       1 NRQRDPQQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQKRYEEQQREDEEKY 60
           |||||||||||||||:|||| |||||||| ||||||||||||||||||||||||||||||
Db     117 NRQRDPQQQYEQCQERCQRHETEPRHMQTCQQRCERRYEKEKRKQQKRYEEQQREDEEKY 176

Qy      61 EERMKEGDN 69
           ||||||  ||
Db     177 EERMKEEDN 185


RESULT   4
Q43358
ID   Q43358      PRELIMINARY;      PRT;   525 AA.
AC   Q43358;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   VICILIN PRECURSOR.
GN   CSV.
OS   Theobroma cacao (Cacao).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids II;
OC   Malvales; Malvaceae; Theobroma.
OX   NCBI_TaxID=3641;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   TISSUE=LEAVES;
RX   MEDLINE=92288309; PubMed=1600151;
RA   McHenry L., Fritz P.J.;
RT   "Comparison of the structure and nucleotide sequences of vicilin genes
RT   of cocoa and cotton raise questions about vicilin evolution.";
RL   Plant Mol. Biol. 18:1173-1176(1992).
DR   EMBL; X62625; CAA44493.1; -.
DR   EMBL; X62626; CAA44494.1; -.
DR   HSSP; P02853; 2PHL.
DR   MENDEL; 30919; Thecc;1188;30919.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
DR   PRODOM; PD081059; -; 1.
KW   Signal.
FT   SIGNAL        1     24       POTENTIAL.
FT   CHAIN        25    525       VICILIN.
SQ   SEQUENCE   525 AA;  60798 MW;  19114CD5C248905D CRC64;


  Query Match             45.4%;  Score 173;  DB 10;  Length 525;
  Best Local Similarity   32.6%;  Pred. No. 1.6e-08;
  Matches   31;  Conservative   17;  Mismatches   15;  Indels   32;  Gaps    1;

Qy       3 QRDPQQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQ---------------- 46
           :|||:|||||||:||:   || |  : |:|||| |::::|:|:
Db      35 ERDPRQQYEQCQRRCESEATEEREQEQCEQRCEREYKEQQRQQEEELQRQYQQCQGRCQE 94

Qy      47 ----------------KRYEEQQREDEEKYEERMK 65
                           ::|:|||:| : | |    |
Db      95 QQQGQREQQQCQRKCWEQYKEQERGEHENYHNHKK 129


RESULT   5
Q9SEW4
ID   Q9SEW4      PRELIMINARY;      PRT;   593 AA.
AC   Q9SEW4;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   VICILIN-LIKE PROTEIN PRECURSOR (FRAGMENT).
OS   Juglans regia (English walnut).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Fagales; Juglandaceae; Juglans.
OX   NCBI_TaxID=51240;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=CV. SUNLAND; TISSUE=SOMATIC EMBRYO LINE;
RA   Teuber S.S., Jarvis K.C., Peterson W.R., Dandekar A.M., Ansari A.A.;
RT   "Identification and cloning of a cDNA encoding a vicilin-like protein,
RT   Jug r 2, from English walnut kernel (Juglans regia): a major food
```

```
RT   allergen.";
RL   Submitted (MAY-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF066055; AAF18269.1; -.
DR   HSSP; P02853; 2PHL.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
FT   NON_TER       1      1
SQ   SEQUENCE   593 AA;  69990 MW;  9BA127E19B18C0D8 CRC64;


  Query Match             43.6%;  Score 166;  DB 10;  Length 593;
  Best Local Similarity   38.4%;  Pred. No. 7.7e-08;
  Matches   38;  Conservative   11;  Mismatches   16;  Indels   34;  Gaps    3;

Qy       2 RQR-----DPQQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKE----------KRKQ 45
           |||     |||||| :||:||| :|  |   : |||||||:|:::          :|:
Db     110 RQRGRDRQDPQQQYHRCQRRCQIQEQSPERQRQCQQRCERQYKEQQGRERGPEASPRRES 169

Qy      46 QKRYEEQQR-----------------EDEEKYEERMKE 66
           : | |||||                 | | || ||  |
Db     170 RGREEEQQRHNPYYFHSQSIRSRHESEEGEVKYLERFTE 208


RESULT   6
Q9ZWI3
ID   Q9ZWI3      PRELIMINARY;      PRT;   810 AA.
AC   Q9ZWI3;
DT   01-MAY-1999 (TrEMBLrel. 10, Created)
DT   01-MAY-1999 (TrEMBLrel. 10, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   PV100.
OS   Cucurbita maxima (Pumpkin) (Winter squash).
OC   Eukaryota; Viridiplantae; Embryophyta; Tracheophyta; Spermatophyta;
OC   Magnoliophyta; eudicotyledons; core eudicots; Rosidae; eurosids I;
OC   Cucurbitales; Cucurbitaceae; Cucurbita.
OX   NCBI_TaxID=3661;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=KUROKAWA AMAKURI NANKIN; TISSUE=COTYLEDON;
RX   MEDLINE=99107919; PubMed=9891029;
RA   Yamada K., Shimada T., Kondo M., Nishimura M., Hara-Nishimura I.;
RT   "Multiple functional proteins are produced by cleaving Asn-Gln bonds
RT   of a single precursor by vacuolar processing enzyme.";
RL   J. Biol. Chem. 274:2563-2570(1999).
DR   EMBL; AB019195; BAA34056.1; -.
DR   HSSP; P02853; 2PHL.
DR   INTERPRO; IPR001113; -.
DR   PFAM; PF00546; Seedstore_7s; 1.
DR   PRODOM; PD081059; -; 1.
SQ   SEQUENCE   810 AA;  97314 MW;  A829A3F7542266AB CRC64;


  Query Match             30.2%;  Score 115;  DB 10;  Length 810;
  Best Local Similarity   32.1%;  Pred. No. 0.004;
  Matches   25;  Conservative   13;  Mismatches   28;  Indels   12;  Gaps    1;

Qy       1 NRQRDPQQQYEQCQKRCQRRETEPRHMQICQQRCERRY-----------EKEKRKQQKR 48
           |::  |: :|| |: |||  |      : |:| || |             | |:|  :
Db      69 NQRGSPRAEYEVCRLRCQVAERGVEQQRKCEQVCEERLREREQGRGEDVDEVERRDPEWE 128

Qy      49 YEEQQREDEEKYEERMKE 66
             |||:| : |: | | :|
Db     129 REEQRRREHEREERRRRE 146


RESULT   7
O60983
ID   O60983      PRELIMINARY;      PRT;  1038 AA.
AC   O60983;
DT   01-AUG-1998 (TrEMBLrel. 07, Created)
DT   01-AUG-1998 (TrEMBLrel. 07, Last sequence update)
DT   01-MAY-2000 (TrEMBLrel. 13, Last annotation update)
DE   REP-LIKE.
GN   REP.
OS   Dictyostelium discoideum (Slime mold).
OG   Plasmid Ddp5.
OC   Eukaryota; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WS2162;
RX   MEDLINE=98198836; PubMed=9539429;
RA   Rieben W.K. Jr., Gonzales C.M., Gonzales S.T., Pilkington K.J.,
RA   Kiyosawa H., Hughes J.E., Welker D.L.;
RT   "Dictyostelium discoideum nuclear plasmid Ddp5 is a chimera related to
RT   the Ddp1 and Ddp2 plasmid families.";
RL   Genetics 148:1117-1125(1998).
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=WS2162;
RA   Rieben W.K., Gonzales C., Gonzales S.T., Pilkington K., Kiyosawa H.,
RA   Hughes J.E., Welker D.L.;
RL   Submitted (APR-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF000580; AAC14390.1; -.
KW   Plasmid.
SQ   SEQUENCE   1038 AA;  121723 MW;  EF24CD47F45BB488 CRC64;


  Query Match             30.1%;  Score 114.5;  DB 5;  Length 1038;
  Best Local Similarity   36.9%;  Pred. No. 0.0055;
  Matches   24;  Conservative   23;  Mismatches   17;  Indels    1;  Gaps    1;

Qy       3 QRDPQQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQKRYEEQ-QREDEEKYE 61
           |:: |:| || ::| |: : | :  | . |:: |:: ::|:::||:: |:| ||| ||: |
Db     936 QQEQQEQQEQRERREQQEQQEQQEQQEQQEQQEQQEQQEQQEQQEQQEQREQEEQEE 995

Qy      62 ERMKE 66
           :| :|
Db     996 QREQE 1000


RESULT   8
Q9VPS3
ID   Q9VPS3      PRELIMINARY;      PRT;   554 AA.
AC   Q9VPS3;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   CG2839 PROTEIN (FRAGMENT).
GN   CG2839.
OS   Drosophila melanogaster (Fruit fly).
OC   Eukaryota; Metazoa; Arthropoda; Tracheata; Hexapoda; Insecta;
OC   Pterygota; Neoptera; Endopterygota; Diptera; Brachycera; Muscomorpha;
OC   Ephydroidea; Drosophilidae; Drosophila.
OX   NCBI_TaxID=7227;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=BERKELEY;
RX   MEDLINE=20196006; PubMed=10731132;
RA   Adams M.D., Celniker S.E., Holt R.A., Evans C.A., Gocayne J.D.,
RA   Amanatides P.G., Scherer S.E., Li P.W., Hoskins R.A., Galle R.F.,
RA   George R.A., Lewis S.E., Richards S., Ashburner M., Henderson S.N.,
RA   Sutton G.G., Wortman J.R., Yandell M.D., Zhang Q., Chen L.X.,
RA   Brandon R.C., Rogers Y.-H.C., Blazej R.G., Champe M., Pfeiffer B.D.,
RA   Wan K.H., Doyle C., Baxter E.G., Helt G., Nelson C.R., Miklos G.L.G.,
RA   Abril J.F., Agbayani A., An H.-J., Andrews-Pfannkoch C., Baldwin D.,
RA   Ballew R.M., Basu A., Baxendale J., Bayraktaroglu L., Beasley E.M.,
RA   Beeson K.Y., Benos P.V., Berman B.P., Bhandari D., Bolshakov S.,
RA   Borkova D., Botchan M.R., Bouck J., Brokstein P., Brottier P.,
RA   Burtis K.C., Busam D.A., Butler H., Cadieu E., Center A., Chandra I.,
RA   Cherry J.M., Cawley S., Dahlke C., Davenport L.B., Davies P.,
RA   de Pablos B., Delcher A., Deng Z., Mays A.D., Dew I., Dietz S.M.,
RA   Dodson K., Doup L.E., Downes M., Dugan-Rocha S., Dunkov B.C., Dunn P.,
RA   Durbin K.J., Evangelista C.C., Ferraz C., Ferriera S., Fleischmann W.,
RA   Fosler C., Gabrielian A.E., Garg N.S., Gelbart W.M., Glasser K.,
RA   Glodek A., Gong F., Gorrell J.H., Gu Z., Guan P., Harris M.,
```

```
RA   Harris N.L., Harvey D., Heiman T.J., Hernandez J.R., Houck J.,
RA   Hostin D., Houston K.A., Howland T.J., Wei M.-H., Ibegwam C.,
RA   Jalali M., Kalush F., Karpen G.H., Ke Z., Kennison J.A., Ketchum K.A.,
RA   Kimmel B.E., Kodira C.D., Kraft C., Kravitz S., Kulp D., Lai Z.,
RA   Lasko P., Lei Y., Levitsky A.A., Li J., Li Z., Liang Y., Lin X.,
RA   Liu X., Mattei B., McIntosh T.C., McLeod M.P., McPherson D.,
RA   Merkulov G., Milshina N.V., Mobarry C., Morris J., Moshrefi A.,
RA   Mount S.M., Moy M., Murphy B., Murphy L., Muzny D.M., Nelson D.L.,
RA   Nelson D.R., Nelson K.A., Nixon K., Nusskern D.R., Pacleb J.M.,
RA   Palazzolo M., Pittman G.S., Pan S., Pollard J., Puri V., Reese M.G.,
RA   Reinert K., Remington K., Saunders R.D.C., Scheeler F., Shen H.,
RA   Shue B.C., Siden-Kiamos I., Simpson M., Skupski M.P., Smith T.,
RA   Spier E., Spradling A.C., Stapleton M., Strong R., Sun E.,
RA   Svirskas R., Tector C., Turner R., Venter E., Wang A.H., Wang X.,
RA   Wang Z.-Y., Wassarman D.A., Weinstock G.M., Weissenbach J.,
RA   Williams S.M., Woodage T., Worley K.C., Wu D., Yang S., Yao Q.A.,
RA   Ye J., Yeh R.-F., Zaveri J.S., Zhan M., Zhang G., Zhao Q., Zheng L.,
RA   Zheng X.H., Zhong F.N., Zhong W., Zhou X., Zhu S., Zhu X., Smith H.O.,
RA   Gibbs R.A., Myers E.W., Rubin G.M., Venter J.C.;
RT   "The genome sequence of Drosophila melanogaster.";
RL   Science 287:2185-2195(2000).
DR   EMBL; AE003588; AAF51469.1; -.
DR   FLYBASE; FBgn0031273; CG2839.
DR   INTERPRO; IPR001304; -.
DR   PFAM; PF00059; lectin_c; 1.
DR   PROSITE; PS50041; C_TYPE_LECTIN_2; 1.
FT   NON_TER     554     554
SQ   SEQUENCE   554 AA;  69873 MW;  029B6D852609FD16 CRC64;


  Query Match             28.6%;  Score 109;  DB 5;  Length 554;
  Best Local Similarity   32.3%;  Pred. No. 0.01;
  Matches   21;  Conservative   24;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 RQRDPQQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQKRYEEQQREDEEKYE 61
           :::: ::: |: :|  :||: | |  :  ::  ||| || :| :::| ||::|::||: |
Db     378 KRKEEERRKEEERKEEERRKEEERKEEERRKEEERRKEKRRRDEKRRREEEKRKEEERKE 437

Qy      62 ERMKE 66
           |  :|
Db     438 EERRE 442


RESULT   9
O40947
ID   O40947      PRELIMINARY;      PRT;  1089 AA.
AC   O40947;
DT   01-JAN-1998 (TrEMBLrel. 05, Created)
DT   01-JAN-1998 (TrEMBLrel. 05, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   ORF 73.
OS   Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Rhadinovirus.
OX   NCBI_TaxID=37296;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97296220; PubMed=9151804;
RA   Neipel F., Albrecht J.C., Fleckenstein B.;
RT   "Cell-homologous genes in the Kaposi's sarcoma-associated rhadinovirus
RT   human herpesvirus 8: determinants of its pathogenicity?";
RL   J. Virol. 71:4187-4192(1997).
DR   EMBL; U93872; AAB62657.1; -.
DR   INTERPRO; IPR002017; -.
SQ   SEQUENCE   1089 AA;  126232 MW;  48AE2F9837E1AB6A CRC64;


  Query Match             28.6%;  Score 109;  DB 12;  Length 1089;
  Best Local Similarity   32.3%;  Pred. No. 0.018;
  Matches   21;  Conservative   24;  Mismatches   20;  Indels    0;  Gaps    0;

Qy       2 RQRDPQQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQKRYEEQQREDEEKYE 61
           :|:| ||| || |   :::: : :  |  |   |:: |:|::::|:: |||::| :|: |
Db     589 QQQDEQQQDEQQQDEQEQQDEQQQDEQQQQDEQEQQEEQEQQEEQEQQEEQEQEQQEEQE 648

Qy      62 ERMKE 66
           : ::|
Db     649 QELEE 653


RESULT  10
Q17909
ID   Q17909      PRELIMINARY;      PRT;   385 AA.
AC   Q17909;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   C11G6.3 PROTEIN.
GN   C11G6.3.
OS   Caenorhabditis elegans.
OC   Eukaryota; Metazoa; Nematoda; Chromadorea; Rhabditida; Rhabditoidea;
OC   Rhabditidae; Peloderinae; Caenorhabditis.
OX   NCBI_TaxID=6239;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Wilkinson J.;
RL   Submitted (MAR-1996) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=94150718; PubMed=7906398;
RA   Wilson R., Ainscough R., Anderson K., Baynes C., Berks M.,
RA   Bonfield J., Burton J., Connell M., Copsey T., Cooper J., Coulson A.,
RA   Craxton M., Dear S., Du Z., Durbin R., Favello A., Fulton L.,
RA   Gardner A., Green P., Hawkins T., Hillier L., Jier M., Johnston L.,
RA   Jones M., Kershaw J., Kirsten J., Laister N., Latreille P.,
RA   Lightning J., Lloyd C., Mcmurray A., Mortimore B., O'Callaghan M.,
RA   Parsons J., Percy C., Rifken L., Roopra A., Saunders D., Shownkeen R.,
RA   Smaldon N., Smith A., Sonnhammer E., Staden R., Sulston J.,
RA   Thierry-Mieg J., Thomas K., Vaudin M., Vaughan K., Waterston R.,
RA   Watson A., Weinstock L., Wilkinson-Sproat J., Wohldman P.;
RT   "2.2 Mb of contiguous nucleotide sequence from chromosome III of C.
RT   elegans.";
RL   Nature 368:32-38(1994).
DR   EMBL; Z70204; CAA94113.1; -.
DR   INTERPRO; IPR001965; -.
DR   PFAM; PF00628; PHD; 1.
SQ   SEQUENCE   385 AA;  46052 MW;  D080FC663B47EDC1 CRC64;


  Query Match             27.7%;  Score 105.5;  DB 5;  Length 385;
  Best Local Similarity   28.4%;  Pred. No. 0.016;
  Matches   21;  Conservative   24;  Mismatches   20;  Indels    9;  Gaps    1;

Qy       2 RQRDPQQQYEQCQKRCQRRETEPRHMQ---------ICQQRCERRYEKEKRKQQKRYEEQ 52
           |:|| ::: |: ||  :|' :   | ::         : ::  ||| |||||:::|: :::
Db     146 RERDERKERERQQKEKEREDAARREIEEKAEMDAKRVAEEEEERRKEKEKRREEKKKQKE 205

Qy      53 QREDEEKYEERMKE 66
             :::|: | : ||
Db     206 LLKEKERSERKEKE 219


RESULT  11
Q98148
ID   Q98148      PRELIMINARY;      PRT;  1162 AA.
AC   Q98148;
DT   01-FEB-1997 (TrEMBLrel. 02, Created)
DT   01-FEB-1997 (TrEMBLrel. 02, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   ORF73 HOMOLOG.
OS   Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Rhadinovirus.
OX   NCBI_TaxID=37296;
RN   [1]
RP   SEQUENCE FROM N.A.
```

```
RX   MEDLINE=97048116; PubMed=8892957;
RA   Cesarman E., Nador R.G., Bai F., Bohenzky R.A., Russo J.J.,
RA   Moore P.S., Chang Y., Knowles D.M.;
RT   "Kaposi's sarcoma-associated herpesvirus contains G protein-coupled
RT   receptor and cyclin D homologs which are expressed in Kaposi's sarcoma
RT   and malignant lymphoma.";
RL   J. Virol. 70:8218-8223(1996).
RN   [2]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97094384; PubMed=8939871;
RA   Moore P.S., Bashoff C., Weiss R.A., Chang Y.;
RT   "Molecular mimicry of human cytokine and cytokine response pathway
RT   genes by KSHV.";
RL   Science 274:1739-1744(1996).
RN   [3]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=97121480; PubMed=8962146;
RA   Russo J.J., Bohenzky R.A., Chien M.C., Chen J., Yan M., Maddalena D.,
RA   Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.;
RT   "Nucleotide sequence of the Kaposi sarcoma-associated herpesvirus
RT   (HHV8).";
RL   Proc. Natl. Acad. Sci. U.S.A. 93:14862-14867(1996).
RN   [4]
RP   SEQUENCE FROM N.A.
RA   Russo J.J., Bohenzky R.A., Chien M.C., Chen J., Yan M., Maddalena D.,
RA   Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.;
RL   Submitted (OCT-1996) to the EMBL/GenBank/DDBJ databases.
RN   [5]
RP   SEQUENCE FROM N.A.
RA   Russo J.J., Bohenzky R.A., Chien M.C., Chen J., Yan M., Maddalena D.,
RA   Parry J.P., Peruzzi D., Edelman I.S., Chang Y., Moore P.S.;
RL   Submitted (MAY-1997) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; U52064; AAC55944.1; -.
DR   EMBL; U75698; AAC57158.1; -.
DR   INTERPRO; IPR002017; -.
SQ   SEQUENCE   1162 AA;  135213 MW;  1A72CE01C1CB081C CRC64;


  Query Match             26.8%;  Score 102;  DB 12;  Length 1162;
  Best Local Similarity   31.3%;  Pred. No. 0.081;
  Matches   21;  Conservative   24;  Mismatches   22;  Indels    0;  Gaps    0;

Qy       2 RQRDPQQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQKRYEEQQREDEEKYE 61
           :|:| ||| || |   |: | | :  |  |   |:: |::: :||:: |:||:::::: :
Db     680 QQQDEQQQDEQQQDEQQQDEQEQQDEQEQQDEQEQQDEQQQDEQQQQDEQQQQDEQQQQD 739

Qy      62 ERMKEGD 68
           |: :: :
Db     740 EQQQQDE 746


RESULT  12
Q9TW28
ID   Q9TW28      PRELIMINARY;      PRT;  1737 AA.
AC   Q9TW28;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-OCT-2000 (TrEMBLrel. 15, Last annotation update)
DE   MYOM PROTEIN.
GN   MYOM.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX2;
RA   Oishi N., Adachi H., Sutoh K.;
RT   "Dictyostelium discoideum gene for MyoM.";
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
RN   [2]
RP   SEQUENCE FROM N.A.
RC   STRAIN=AX2;
RA   Geissler H., Schwarz E.C., Soldati T.;
RT   "Identification of two novel and highly divergent myosins in
RT   Dictyostelium discoideum.";
RL   Submitted (SEP-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AB017910; BAA84604.1; -.
DR   EMBL; AF090533; AAD47903.1; -.
DR   HSSP; P08799; 1MND.
DR   INTERPRO; IPR000194; -.
DR   INTERPRO; IPR000219; -.
DR   INTERPRO; IPR001609; -.
DR   INTERPRO; IPR001849; -.
DR   PFAM; PF00063; myosin_head; 2.
DR   PFAM; PF00169; PH; 1.
DR   PFAM; PF00621; RhoGEF; 1.
DR   PRINTS; PR00193; MYOSINHEAVY.
DR   PROSITE; PS00152; ATPASE_ALPHA_BETA; UNKNOWN_1.
DR   PROSITE; PS50003; PH_DOMAIN; 1.
SQ   SEQUENCE   1737 AA;  195860 MW;  39CBF9C08EFEF856 CRC64;


  Query Match             26.5%;  Score 101;  DB 5;  Length 1737;
  Best Local Similarity   35.8%;  Pred. No. 0.14;
  Matches   24;  Conservative   20;  Mismatches   19;  Indels    4;  Gaps    2;

Qy       1 NRQRDPQQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKE-KRKQQKRYEEQQREDEEK 59
             |:::: ::: |      ||:| | |  :: :|| |   | | |||:::|  |:||::|||
Db     940 NKRKEEERKKE---LERQRKEEEERQKELERQRREEEKELERKRKEEERELERQRKEEEK 996

Qy      60 YEERMKE 66
             :|| ::
Db     997 EQERKRK 1003


RESULT  13
Q9Y0C9
ID   Q9Y0C9      PRELIMINARY;      PRT;   838 AA.
AC   Q9Y0C9;
DT   01-NOV-1999 (TrEMBLrel. 12, Created)
DT   01-NOV-1999 (TrEMBLrel. 12, Last sequence update)
DT   01-NOV-1999 (TrEMBLrel. 12, Last annotation update)
DE   RAS INTERACTING PROTEIN RIPA.
GN   RIPA.
OS   Dictyostelium discoideum (Slime mold).
OC   Eukaryota; Dictyosteliida; Dictyostelium.
OX   NCBI_TaxID=44689;
RN   [1]
RP   SEQUENCE FROM N.A.
RA   Lee S., Parent C.A., Insall R., Firtel R.A.;
RT   "The novel ras interacting protein required for chemotaxis and cAMP
RT   signal relay in Dictyostelium.";
RL   Submitted (JUN-1999) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; AF159241; AAD43567.1; -.
SQ   SEQUENCE   838 AA;  94059 MW;  EAAD911294226808 CRC64;


  Query Match             26.2%;  Score 100;  DB 5;  Length 838;
  Best Local Similarity   24.2%;  Pred. No. 0.093;
  Matches   16;  Conservative   33;  Mismatches   17;  Indels    0;  Gaps    0;

Qy       1 NRQRDPQQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQKRYEEQQREDEEKY 60
           |:|:  ||| :| |:: |::: : :| |  ||: ::: ::::::||:: ::||:: :::
Db      88 NQQQQQQQQAQQQQQQQQQQQQQQQHQQQQQQQQQQQQQQQQQQQQQQAQQQQQQAQQQA 147

Qy      61 EERMKE 66
           : : ::
Db     148 QVQAQQ 153


RESULT  14
Q14151
ID   Q14151      PRELIMINARY;      PRT;   953 AA.
AC   Q14151;
DT   01-NOV-1996 (TrEMBLrel. 01, Created)
DT   01-NOV-1996 (TrEMBLrel. 01, Last sequence update)
```

```
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   KIAA0138 PROTEIN.
GN   KIAA0138.
OS   Homo sapiens (Human).
OC   Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
OC   Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
OX   NCBI_TaxID=9606;
RN   [1]
RP   SEQUENCE FROM N.A.
RX   MEDLINE=96127530; PubMed=8590280;
RA   Nagase T., Seki N., Tanaka A., Ishikawa K., Nomura N.;
RT   "Prediction of tbe coding sequences of unidentified human genes. IV.
RT   The coding sequences of 40 new genes (KIAA0121-KIAA0160) deduced by
RT   analysis of cDNA clones from human cell line KG-1.";
RL   DNA Res. 2:167-174(1995).
RN   [2]
RP   SEQUENCE FROM N.A.
RA   Lamerdin J.E., McCready P.M., Skowronski E., Adamson A.W.,
RA   Burkhart-Schultz K., Gordon L., Kyle A., Ramirez M., Stilwagen S.,
RA   Phan H., Velasco N., Garnes J., Danganan L., Poundstone P.,
RA   Christensen M., Georgescu A., Avila J., Liu S., Attix C., Andreise T.,
RA   Trankheim M., Amico-Keller G., Coefield J., Duarte S., Lucas S.,
RA   Bruce R., Thomas P., Quan G., Kronmiller B., Arellano A.,
RA   Montgomery M., Ow D., Nolan M., Trong S., Kobayashi A., Olsen A.O.,
RA   Carrano A.V.;
RL   Submitted (APR-1998) to the EMBL/GenBank/DDBJ databases.
DR   EMBL; D50928; BAA09487.1; -.
DR   EMBL; AC004611; AAC14666.1; -.
DR   INTERPRO; IPR000504; -.
DR   INTERPRO; IPR002106; -.
DR   INTERPRO; IPR003034; -.
DR   PFAM; PF00076; rrm; 1.
DR   PFAM; PF02037; SAP; 1.
DR   PROSITE; PS00179; AA_TRNA_LIGASE_II_1; UNKNOWN_1.
SQ   SEQUENCE   953 AA;  107473 MW;  084343934F8B3196 CRC64;


  Query Match             26.2%;  Score 100;  DB 4;  Length 953;
  Best Local Similarity   35.1%;  Pred. No. 0.1;
  Matches   26;  Conservative   17;  Mismatches   21;  Indels   10;  Gaps    4;

Qy       2 RQRDPQQQYEQCQKRCQRRETEPRHMQICQQRCER---RYEKEKRKQQKRY----EEQQR 54
           |:    : | |: |  |||:  |    |:  ::| ||   | |:|:||:|:|     || :|
Db     657 RKEKARLQRERLQLECQRQRLERERME--RERLERERMRVERERRKEQERIHREREELRR 714

Qy      55 EDEE-KYEERMKEG 67
           : |: :||:  : |
Db     715 QQEQLRYEQERRPG 728


RESULT  15
Q9QR71
ID   Q9QR71      PRELIMINARY;      PRT;  1129 AA.
AC   Q9QR71;
DT   01-MAY-2000 (TrEMBLrel. 13, Created)
DT   01-MAY-2000 (TrEMBLrel. 13, Last sequence update)
DT   01-JUN-2000 (TrEMBLrel. 14, Last annotation update)
DE   LATENT NUCLEAR ANTIGEN.
GN   ORF 73.
OS   Kaposi's sarcoma-associated herpesvirus (KSHV) (Human herpesvirus 8).
OC   Viruses; dsDNA viruses, no RNA stage; Herpesviridae;
OC   Gammaherpesvirinae; Rhadinovirus.
OX   NCBI_TaxID=37296;
RN   [1]
RP   SEQUENCE FROM N.A.
RC   STRAIN=GK18;
RA   Glenn M., Rainbow L., Aurade F., Davison A., Schulz T.F.;
RT   "Identification of a Spliced Gene from Kaposi's Sarcoma-Associated
RT   Herpesvirus Encoding a Protein with Similarities to Latent Membrane
RT   Proteins 1 and 2A of Epstein-Barr Virus.";
RL   J. Virol. 73:6953-6963(1999).
DR   EMBL; AF148805; AAD46501.1; -.
DR   INTERPRO; IPR000533; -.
DR   INTERPRO; IPR002017; -.
DR   INTERPRO; IPR002965; -.
DR   PRINTS; PR00194; TROPOMYOSIN.
DR   PRINTS; PR01217; PRICHEXTENSN.
SQ   SEQUENCE   1129 AA;  131346 MW;  8F63855B45F79109 CRC64;


  Query Match             26.2%;  Score 100;  DB 12;  Length 1129;
  Best Local Similarity   34.3%;  Pred. No. 0.12;
  Matches   23;  Conservative   21;  Mismatches   21;  Indels    2;  Gaps    1;

Qy       2 RQRDPQQQYEQCQKRCQRRETEPRHMQICQQRCERRYEKEKRKQQKRYE--EQQREDEEK 59
           :|:| ||| || |   :::: : :  |  |   |:: |:|::::|:: |  ||| | |::
Db     666 QQQDEQQQDEQQQDEQEQQDEQQQDEQQQQDEQEQQEEQEQQEEQEQQEEQEQQEEQE 725

Qy      60 YEERMKE 66
             ||: :|
Db     726 LEEQEQE 732


Search completed: March  1, 2001, 16:09:25
Job time: 1569 sec
```